Lesson 1.3 - February 4, 2021

Review

- Yes and No questions will have 0 associated with "No" and 1 associated with "Yes" (Binary System)
- Recall the difference between parameters and statistics
- Recall the difference between stacked and unstacked data sets

---

Ways to Organize Categorical Data

- Something to keep in mind: when would you want to modify the counts or the numbers included in the data or change the actual numbers into percentages or proportions
- Relative Frequency can provide key information when studying a specific dataset

---

Mathematical Standards

- When calculating a fraction in statistics, the denominator should always be the total
- A specific percentage will be its respective fraction multiplied by 100

Observational Studies vs. Controlled Experiment

- Observational Studies are used to show associations
  - Can be aggected by confounding variables (variables that may affect final results)
- Controlled Experiments are used to prove cause and effects
  - Controlled experiment tries to minimize lurking and confounding variables
- In both, the goal is to find how one variable affects another variable
- For both, you want to split the participants into two groups as evenly as possible (treatment group and control group)
- In an obsevational study, the treatment and control groups are chosen randomly
- In controlled experiments, you can control who is in the treatment and control groups

- In certain cases, it is impossible to turn the event into a controlled experiment (smoker example)
- The two groups in an eperiment should be as similar as possible and the only difference between the people in the two groups should be the treatment variable
- **Confounding Variable:** A variable other than the treatment variable that causes a change in the result

---

Controlled Experiment Standards

- The sample size must be large enough to have enough opportunity to study varying cases and have higher accuracy
- The study should use a placebo if possible
- The subjects should be placed in treatment and control group at random
- Ideally, the study should be "double-blind" - people in the experiment shouldn't know whether they are in the treatment or control group and the people administering the treatment do not know this quality either